

EMERGENCE OF FUNCTION

INDIKA RAJAPAKSE AND STEVE SMALE

ABSTRACT. This work gives a mathematical study of tissue dynamics. We combine within-cell genome dynamics and diffusion between cells, where the synthesis of the two gives rise to the emergence of function. We introduce a concept of monotonicity and prove that monotonicity together with hardwiring, defined as all cells of the same tissue having the same genome dynamics, is sufficient for the global convergence of the tissue dynamics.

One of the most beautiful questions in biology is how individual cells and tissues, each expressing information from a single genome, give rise to all functions in a multicellular organism. Is there a basis for the emergence of tissue-specific function? In vertebrates, consider the liver, functioning to detoxify and ensure an appropriate composition of blood, and the skeletal muscle, functioning to contract and generate force. In each of these tissues millions of individual cells contribute to emergence of function according to their cell type.

Tissue here means a set of cells of the same cell type located together as exemplified by an organ in the body.

Understanding proteins is central to understanding this emergence from single cells to a whole tissues. The main elements of emergence that we consider are first, the unique protein distribution in a given cell type and second, the cellular architecture of the tissue, a three dimensional structure with diffusion of molecules between cells.

We build a mathematical model for emergence of function, drawing on our previous work on cell dynamics (genome dynamics), and the work of Alan Turing on diffusion [1]. Here we combine the cell dynamics and the diffusion between cells, where the synthesis of the two gives rise to the emergence of function. Conditions are investigated under which the dynamics of the tissue of an organism converge to an equilibrium where the proteins of individual cells have the same distribution. Underlying our setting are known biological phenomena: 1) cells within a tissue (i.e. the same cell type) have the same dynamics and distribution of proteins at equilibrium 2) the function of a cell corresponds to the proteins of that cell. For reasons to be discussed, we call the property in 1) "hardwiring" of the tissue [2]. Convergence of the tissue dynamics to such an equilibrium naturally takes on importance, for its role in maintenance of tissue function. Even a local stability of the (hardwiring) equilibrium, i.e. it's robustness, gives some validity to our model in biology. We introduce a property of cell dynamics (first for one cell and then extended to many cells) that we call monotonicity. Our main theorem (Theorem 5) establishes that monotonicity implies global convergence of the tissue dynamics

We extend thanks to James Gimlett and Srikanta Kumar at Defense Advanced Research Projects Agency for support and encouragement.

to the equilibrium, where all cells have the same protein distribution. This gives a biological justification for our framework, and a model for "emergence of function," as well as suggestions for studying the passage from emergence to morphogenesis.

We cover the following content in this work:

1. A simple example
2. One cell and one protein from the gradient point of view
3. The genome dynamics of one cell and n proteins
4. Cellular dynamics with a single protein
5. Dynamics of a tissue (m cells and n proteins)
6. Turing's paper on morphogenesis
7. Lapse of emergence

1. Simple example

Here we model two cells, separated by a membrane, that each have a same single protein. Consider the following system,

$$(1a) \quad \text{Cell 1:} \quad \frac{dx}{dt} = a(x - x_0)$$

$$(1b) \quad \text{Cell 2:} \quad \frac{dy}{dt} = b(y - y_0), \quad a, b, < 0, x_0, y_0 > 0,$$

where x and y can be interpreted as protein concentration, x in cell 1 and y cell 2, both positive. Thus, $x, y \in \mathbf{X} \times \mathbf{Y} = [0, c] \times [0, d]$, where c and d represent the maximum concentration of protein x and y respectively. The equilibria are: $x = x_0$ and $y = y_0$.

We introduce Turing type (diffusion) coupling by adding a term with $\beta > 0$ as follows:

$$(2) \quad \begin{aligned} \frac{dx}{dt} &= a(x - x_0) + \beta(y - x) \\ \frac{dy}{dt} &= b(y - y_0) + \beta(x - y), \quad \beta > 0 \end{aligned}$$

The equilibrium for the above system is obtained by solving the system derived from Equation 2) by setting the right hand sides equal to zero. This is a linear system in two equations and two variables and we obtain

$$(3a) \quad x = \frac{-abx_0 + a\beta x_0 + b\beta y_0}{a\beta + b\beta - ab}$$

$$(3b) \quad y = \frac{-aby_0 + a\beta x_0 + b\beta y_0}{a\beta + b\beta - ab}.$$

It is not hard to see from Equation 3a and 3b that if $\beta \rightarrow \infty$, x and y converge to the same value $\frac{ax_0 + by_0}{a+b}$. Therefore this system approaches a common protein concentration and the example exhibits the role of diffusion, even with different cell dynamics. We refer to this as an "emergent equilibrium." Eigenvalues of the Jacobian matrix (see the following sections) for system 2 at equilibrium and finite

FIGURE 1. Plot of Equations 3a and 3b for $a = -2, b = -1, c = -1, d = -2$ as β changes. x (red) and y (blue) are the coordinates of the equilibrium of Equation 2.

β are expressed as:

$$\begin{aligned} \frac{1}{2}(a+b) - \beta + \frac{1}{2}\sqrt{(a-b)^2 + 4\beta^2} \\ \frac{1}{2}(a+b) - \beta - \frac{1}{2}\sqrt{(a-b)^2 + 4\beta^2}. \end{aligned}$$

Since the eigenvalues are real negative, this pair (x, y) of Equation 3 is a stable equilibrium.

Success of emergence: The magnitude of $(x - y)$ from Equation 3 measures the departure from the "emergence" as,

$$(4) \quad x - y = -\frac{ab(x_0 - y_0)}{\beta(a+b) - ab}.$$

If $x_0 - y_0$ is big and β is small there is ill-conditioning as follows. If $a = 0$ in Equation (3), the solution is $x = y_0$ and $y = y_0$. If $a \neq 0$ no matter how small, and $\beta = 0$, the solution is $x = x_0$ and $y = y_0$.

Note from Equation 4, for any finite β the equilibrium for the pair (x, y) has x not y if $x_0 \neq y_0$. We might say then the system 2 is not emergent (for any finite β).

Figure 1 show a numerical example of this system.

Remark 1: Here the β anticipates the fiedler number of a laplacian defined by the cellular network of the tissue. We will introduce the concept of a "Hardwiring hypothesis", which implies $x_0 = y_0$. Diffusion is unnecessary for emergence (and in fact it can defeat emergence (!) as we will see). On the other hand diffusion can have a stabilizing effect.

2. One cell and one protein from the gradient point of view

We given an alternative point of view of genome dynamics that will not be used in the rest of the paper. This section may therefore be safely skipped. Consider the dynamics described as the gradient of a potential function.

FIGURE 2. The blue curve is the potential function $f(x) = 2(x - x_0)^2 + 1$ and the black curve is the negative gradient of $f(x)$.

Suppose f is quadratic

$$(5) \quad f(x) = \frac{a}{2}(x - x_0)^2 + b, \quad x > 0, x_0 > 0, a, b > 0$$

where the derivative $f'(x) = a(x - x_0)$, $\arg \min f(x) = x_0$, $f(x_0) = b$ and a is the rate of convergence to the equilibrium (see Figure 2). f could be interpreted as a potential function and x as a protein concentration. From Equation 5, we obtain an example of genome dynamics as is in Section 1.

$$(6) \quad \frac{dx}{dt} = -(\text{Gradient } f(x)), \quad x > 0$$

where the derivative $f'(x) = a(x - x_0)$, $\min f(x) = b$. Substituting Equation 6 into Equation 5 we obtain an example of genome dynamics as is in Section 1.

$$(7) \quad \frac{dx}{dt} = -a(x - x_0).$$

The solution to Equation 7 is

$$(8) \quad x(t) = C \exp^{-at} + x_0$$

checked as follows

$$(9) \quad \frac{dx(t)}{dt} = -aC \exp^{-at} = -a(x(t) - x_0).$$

Solving for C by setting $t = 0$ in Equation 9

$$C = x(0) - x_0, \quad x(0) = \text{initial condition of (8)}$$

Therefore

$$x(t) = (x(0) - x_0)e^{-at} + x_0, \quad x_0 > 0, a > 0.$$

The equilibrium for $\frac{dx}{dt} = -a(x - x_0)$ is given by x_0 . Since the second derivative of f in Equation 6 is negative, the equilibrium is a stable equilibrium.

Figure 2 is an example of a potential function and its $-$ gradient.

This is an example of linear dynamics of one protein and one cell with stability. This dynamics while linear is also a good approximation of the general stable dynamics

in a neighborhood of the equilibria. Moreover the global dynamics of the basin of the stable equilibrium is qualitatively equivalent to the linear example.

3. The genome dynamics of one cell and n proteins

We use the setting of our paper on genome dynamics [2]. For a single cell state space $\mathbf{X} = \prod_{j=1}^n [0, c_j]$, where c_j is the maximum concentration of protein j . Sometime we use $c_j = \infty$. The inner product is Cartesian. The genome dynamics is expressed as $\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x})$, where \mathbf{F} is a function from $\mathbf{X} \subset \mathbb{R}^n$ to \mathbb{R}^n .

Generally recall the notion of stable equilibrium \mathbf{x}_0 of \mathbf{X} for $\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x})$, as well as its basin. If every eigenvalue of the Jacobian matrix of first partial derivatives of \mathbf{F} at \mathbf{x}_0 has negative real part, then all trajectories that start near \mathbf{x}_0 approach it as $t \rightarrow \infty$. The basin $\mathbf{B}(\mathbf{x}_0)$ is the set of all points which tends to \mathbf{x}_0 when $t \rightarrow \infty$. Then \mathbf{x}_0 is a stable equilibrium.

The dynamics on the basin is "linear" provided that the $\mathbf{F} : \mathbf{B} \rightarrow \mathbb{R}^n$, $\mathbf{B} \subset \mathbb{R}^n$, has the form $\mathbf{F}(\mathbf{x}) = \mathbf{F}_{lin}(\mathbf{x} - \mathbf{x}_0)$, where \mathbf{F}_{lin} is a linear map, $\mathbb{R}^n \rightarrow \mathbb{R}^n$, \mathbf{x} belongs to \mathbf{B} and \mathbf{x}_0 is the equilibrium in \mathbf{B} .

Suppose that \mathbf{F} is the dynamics, not necessarily linear, with stable equilibrium \mathbf{x}_0 . Then \mathbf{F} at \mathbf{x}_0 has the above form in a neighborhood of \mathbf{x}_0 (\mathbf{F}_{lin} is the derivative of \mathbf{F} at \mathbf{x}_0). More over it can be shown that the dynamics on the basin is topologically equivalent to the linear dynamics as above.

In fact a main theorem about stable equilibria is that the linear dynamic in \mathbb{R}^n is equivalent to the dynamics of $\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x})$ in the basin $\mathbf{B}(\mathbf{x}_0)$. This means there is a homeomorphism from \mathbb{R}^n to the basin $\mathbf{B}(\mathbf{x}_0)$ that preserves the solution curves. Then the dynamics of $\mathbf{B}(\mathbf{x}_0)$ is same topologically as linear dynamics above.

Browder [3, 4] and Hirsch [5, 6] have extensively studied the topic of monotonicity. Here we provide our version of monotonicity which is new, has a common element with Hirsch and Browder, but is quite different.

Definition 1: Monotonicity condition. Suppose we have a dynamics $\frac{d\mathbf{x}}{dt} = \mathbf{G}(\mathbf{x})$ on a domain \mathbf{X} in a euclidean space with it's inner product. The monotonicity condition for \mathbf{G} , and a point $\mathbf{x}_0 \in \mathbf{X}$ is then:

$$(10) \quad \langle \mathbf{G}\mathbf{x}, (\mathbf{x} - \mathbf{x}_0) \rangle < 0 \text{ for all } \mathbf{x} \neq \mathbf{x}_0 \text{ in } \mathbf{X}.$$

Proposition 1: For any dynamics $\frac{d\mathbf{x}}{dt} = \mathbf{G}(\mathbf{x})$ on $\mathbf{X} \subset \mathbb{R}^k$, $\mathbf{x}_0 \in \mathbf{X}$, the monotonicity condition for \mathbf{X} , \mathbf{x}_0 , implies that $\|\mathbf{x}(t) - \mathbf{x}_0\|$ is a decreasing function of t for all non trivial solutions $\mathbf{x}(t)$ in \mathbf{X} , where $\mathbf{x}(t)$ defined for all $t > 0$.

Proof: Suppose Equation 10 is true. Note that

$$\begin{aligned} \frac{d(\|\mathbf{x}(t) - \mathbf{x}_0\|^2)}{dt} &= \frac{d\langle \mathbf{x}(t) - \mathbf{x}_0, \mathbf{x}(t) - \mathbf{x}_0 \rangle}{dt} \\ &= 2 \left\langle \frac{d(\mathbf{x}(t))}{dt}, \mathbf{x}(t) - \mathbf{x}_0 \right\rangle = 2 \langle \mathbf{G}\mathbf{x}(t), (\mathbf{x}(t) - \mathbf{x}_0) \rangle \end{aligned}$$

The quantity at the end is negative by Equation 10, the monotonicity condition. QED

FIGURE 3. $\langle \mathbf{F}p, (p - \mathbf{x}_0) \rangle > 0$. This an example of a basin which is not a monotonic basin.

One could call the \mathbf{X} of Proposition 1, a "monotonic basin" for the dynamics. Under these conditions \mathbf{x}_0 is an equilibrium.

Thus, monotonicity on \mathbf{X}, \mathbf{x}_0 implies that $\mathbf{x}(t)$ is monotonically converging to \mathbf{x}_0 . This gives a stability of \mathbf{x}_0 . The converse is not true not even in the linear case. One can take for an example a spiral sink where the axes are different (Figure 3). When the solution is going in the direction of the long axis then $\mathbf{x}(t) - \mathbf{x}_0$ is not decreasing, while \mathbf{x}_0 is a stable equilibrium. This example helps understand the famous Turing phenomenon (see Section 6).

Example 1: Let $\mathbf{F}\mathbf{x} = \mathbf{A}(\mathbf{x} - \mathbf{x}_0)$, where \mathbf{A} is a linear map $\mathbb{R}^n \rightarrow \mathbb{R}^n$, not necessarily symmetric. Then \mathbf{A} is negative definite exactly when monotonicity holds.

Let us return to the biological setting. Single cell dynamics is that of dynamics on a basin $\mathbf{B} \subset \mathbf{X}$ as in our previous work on genome dynamics [2]. We assume that the basin \mathbf{B} is that of an equilibrium \mathbf{x}_0 and are excluding periodic attractors in the present paper. This means we are identifying a cell with its basin. The equilibrium of a genome dynamics of a cell exhibits the distribution of proteins. That distribution can be identified with that cell.

We now examine explicitly the conditions for monotonicity in the linear case of one cell with two proteins. This case can be represented by the following system represented by the following system

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}\mathbf{x} = \mathbf{A}(\mathbf{x} - \mathbf{x}_0), \quad \mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)$$

where $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $\frac{d\mathbf{x}}{dt} = 0$ when $\mathbf{x} = \mathbf{x}_0$. Then \mathbf{A} (Jacobian matrix at \mathbf{x}_0) is stable and $\mathbf{x}(t) \rightarrow \mathbf{x}_0$ when all eigenvalues have negative real parts. The eigenvalues of \mathbf{A} are given by the characteristic equation $\lambda^2 - \tau\lambda + \Delta = 0$, where

$$\tau = \text{trace}(\mathbf{A}) = a + d \text{ and } \Delta = \det(\mathbf{A}) = ad - bc.$$

Then $\lambda_1 = \frac{\tau + \sqrt{\tau^2 - 4\Delta}}{2}, \lambda_2 = \frac{\tau - \sqrt{\tau^2 - 4\Delta}}{2}$ are the eigenvalues of \mathbf{A} . For stability \mathbf{A} must satisfy two criteria: 1) The trace, $a + d$, must be negative, and 2) the determinant, $ad - bc$, must be positive [6, 7].

To derive the conditions for monotonicity, consider the quadratic form associated with \mathbf{A} : $Q(u, v) = au^2 + dv^2 - \left(\frac{b+c}{2}\right)uv$, and suppose $a, d < 0$. Thus,

FIGURE 4. Monotonicity and stability conditions in the bc plane for $a, d = -1$. The dark shading region together with the blue region (**E**) constitutes the stability region. **E** is the monotonic region. The red dot shows Turing's two-cell example in the bc plane.

$\frac{\mathbf{A} + \mathbf{A}^T}{2} = \begin{pmatrix} a & \frac{b+c}{2} \\ \frac{b+c}{2} & d \end{pmatrix}$, and $\frac{\mathbf{A} + \mathbf{A}^T}{2}$ is symmetric. The matrix of a quadratic form can always be forced to be symmetric in this way. The condition for monotonicity is $\langle \mathbf{F}\mathbf{x}, (\mathbf{x} - \mathbf{x}_0) \rangle < 0$. This amounts to $\langle \mathbf{A}(\mathbf{x} - \mathbf{x}_0), (\mathbf{x} - \mathbf{x}_0) \rangle < 0$ or that the eigenvalues of $\left(\frac{\mathbf{A} + \mathbf{A}^T}{2}\right)$ are negative, which is equivalent to \mathbf{A} being negative definite. Since the determinant of $\frac{\mathbf{A} + \mathbf{A}^T}{2}$ is positive,

$$(11) \quad ad > \left(\frac{b+c}{2}\right)^2, \quad a, d < 0.$$

In summary the stability condition is $bc < ad$ and the monotonicity condition is $\left(\frac{b+c}{2}\right)^2 < ad$. Therefore, excess of the left hand sides of the previous inequalities is $\left(\frac{b+c}{2}\right)^2 - bc$. If the excess is positive or zero, monotonicity implies stability. The excess is never negative.

More generally, as a consequence of Proposition 1 one can prove the following.

Proposition 2: For a linear dynamics on \mathbb{R}^n monotonicity implies stability.

Figure 4 shows an example of monotonicity and stability conditions in the bc plane, where $a, d = -1$. **E** is the monotonic region and hence is part of the stable region. The red dot represents Turing's two-cell example (Turing [1] and Chua [11]), discussed in Section 6.

Hardwiring: The genes present in the human genome are the same in all cell types and all individuals. Now we describe a property of a family of cells, which we called hardwiring [2], motivated by the universality above. Our network in [2] puts an oriented edge (between two nodes), between two genes, i and j , if it is possible for the protein product of gene i to bind to the promoter of gene j and activate transcription. Gene i will bind to this promoter only in some cell types, at certain stages of development. It can happen that gene i as a transcription factor may be

silenced. In that case gene i can be removed from the network together with its edges. As an example, this phenomenon can occur through failure of chromatin accessibility [8, 9]. We will say that a family of cells is hardwired provided that the genome dynamics is the same for every cell in the family. In the example of Turing (also Smale [10], Chua [11, 12]) below hardwiring is assumed extensively.

Definition of weak hardwiring: Thus, the family is hardwired provided that the dynamics of each cell in the family is the same; in particular the equilibrium of each cell is the same. That is, the protein distribution at the equilibrium of each cell is the same. If the last property is true then we will say that the family satisfies "weak hardwiring." The idea of the weak hardwiring concept is that in a single cell type all cells have the same equilibrium distribution of proteins [2]. This helps justify the identification of a tissue with its protein distribution

4. Cellular dynamics and its architecture

We will define a graph \mathbf{G} as a mathematical model for the cellular structure for a single tissue. First consider the m cells of the single tissue and a single protein. The main biological object is the cellular architecture of a tissue which consists of m cells in three dimensions. The graph \mathbf{G} consists of nodes corresponding to the cells of the tissue. The weighted edges of the graph are associated with the membranes between two cells and define the notion of adjacency. This adjacency is represented by a number which represents the diffusion, between two cells and it depends on the interactions at their cell membranes [13]. This number could be interpreted as the product of the permeability and the area of the membrane between cell i and cell j , a quantity represented by a matrix element a_{ij} . We write $\mathbf{A} = (a_{ij})$. The matrix \mathbf{A} is an $m \times m$ symmetric matrix, the adjacency matrix of the architecture. Note that \mathbf{A} does not depend on the protein levels. Thus, \mathbf{G} is a weighted graph whose nodes are $i = 1, \dots, m$, and edges a_{ij} . We assume that the graph is connected. What we have discussed here is a network whose nodes are cells and it not to be confused with the genome network in Section 3. This model applies more literally to diffusion in the case of small molecules.

Definition of a state: A state associated to the graph is a set of protein levels $\mathbf{x}_1, \dots, \mathbf{x}_m$, where \mathbf{x}_i is the level of a single protein in the i^{th} cell. Thus a state \mathbf{x} is a function of nodes i and with value at node i written as \mathbf{x}_i . The states form a linear space \mathbf{S} , and feasible states, the subspace of functions with non-negative values \mathbf{S}_+ . The function $\mathbf{x} \in \mathbf{S}$ is harmonic provided that \mathbf{x}_i is a constant function of i . By our hypothesis that the graph \mathbf{G} is connected, it follows that the space of harmonic functions is one dimensional.

For n proteins we generalize the notion of \mathbf{x}_i and \mathbf{S} . Now \mathbf{x}_i is a distribution of proteins in the i^{th} cell, i. e. $\mathbf{x}_i = (\mathbf{x}_i^1, \dots, \mathbf{x}_i^n)$. Note that this expression can be thought as a function of i . We assume that the membrane structure of the tissue affects all the proteins equally. (this is a strong idealization, but it can be relaxed easily as in Turing's example [1] in Section 6).

The Laplacian matrix: Let \mathbf{D} be the diagonal matrix with the i^{th} element of the diagonal defined by $\sum_j a_{ij}$. Then the Laplacian is given by

$$\mathbf{L} = \mathbf{D} - \mathbf{A}.$$

\mathbf{L} is a $(m \times m)$ real symmetric matrix, and together with the non-negativity of the weights, this implies that it is positive semidefinite [14, 15]. It is an operator on \mathbf{S} .

The diffusion dynamics defined by the cellular architecture may be written as follows:

$$\frac{d\mathbf{x}_i}{dt} = - \sum_{j \in m_i} a_{ij} (\mathbf{x}_i - \mathbf{x}_j), \quad i = 1, \dots, m.$$

or

$$(12) \quad \frac{d\mathbf{x}}{dt} = -\mathbf{L}\mathbf{x}.$$

Note that Equation (12) is a linear system of ordinary differential equations.

Remark 2: Harmonic functions are exactly set of \mathbf{x} , such that $\mathbf{L}\mathbf{x} = \mathbf{0}$.

Note that our definition applies not just to a single protein, but to an n -tuple \mathbf{x}_i belonging to \mathbb{R}^n . The feasible ones to $(\mathbb{R}_+)^n$.

Proposition 3: The system is globally stable with equilibrium set, the harmonic functions.

Proof: $\langle -\mathbf{L}\mathbf{x}, \mathbf{x} \rangle \leq 0$ and $\langle -\mathbf{L}\mathbf{x}, \mathbf{x} \rangle = 0$ iff $\mathbf{x} = \text{constant}$, i.e. $\mathbf{x}_1 = \mathbf{x}_2 = \dots = \mathbf{x}_m$. The solution to Equation (12) will be denoted by $\mathbf{x}(t)$ with initial conditions $\mathbf{x}(0) = \mathbf{C}$, $\mathbf{C} \in \mathbb{R}^m$. Now the solution is

$$\mathbf{x}(t) = e^{-\mathbf{L}t} \mathbf{C}.$$

Then $\frac{d}{dt} \langle \mathbf{x}(t), \mathbf{x}(t) \rangle = 2 \langle -\mathbf{L}\mathbf{x}(t), \mathbf{x}(t) \rangle < 0$, unless $\mathbf{x}(t)$ satisfies $\mathbf{x}_1(t) = \mathbf{x}_2(t) = \dots = \mathbf{x}_m(t) = \text{constant}$. Therefore, the solution converges to a harmonic function. QED

For the n protein case, the harmonic functions form an n -dimensional space defined by $\mathbf{x}_1 = \mathbf{x}_2 = \dots = \mathbf{x}_m$, where \mathbf{x}_1 is an arbitrary element of \mathbb{R}^n .

5. Dynamics of a tissue (m cells and n proteins)

We use both the notations $\mathbf{X}_i = \prod_j^n [0, c_j]$ and the basin, $\mathbf{B}_i \subset \mathbf{X}_i$ with its equilibrium $\mathbf{x}_{0,i}$ for the domain of the genome dynamics, where c_j is the maximum protein concentration protein j and n is number of proteins. \mathbf{X}_i is important for the lapse of emergence and dealing with different cell types (different tissues) as in Section 7. \mathbf{B}_i is suited for single tissue theory as in the following.

Genome dynamics for m cells:

For a single cell say i , \mathbf{B}_i is the domain of the dynamics. For each cell i , $\mathbf{F}_i : \mathbf{B}_i \rightarrow \mathbb{R}^n$ represents the genome dynamics in cell i ,

$$(13) \quad \frac{d\mathbf{x}_i}{dt} = \mathbf{F}_i \mathbf{x}_i, \quad \mathbf{x}_i \in \mathbf{B}_i, \quad i = 1, \dots, m.$$

For the case of cells of a tissue $\mathbf{S} = \prod_i^m \mathbf{B}_i$, $i = 1, \dots, m$, where m is the number of cells in the tissue and \mathbf{S} is the state space of Section 3 extended to n proteins. We use an inner product on \mathbf{S} derived from the inner products on \mathbf{B}_i

Thus \mathbf{B}_i corresponds to cell i with stable equilibrium $\mathbf{x}_{0,i}$.

Now we take the product of the dynamics over all the cells at once to get $\mathbf{F} : \mathbf{S} \rightarrow \mathbb{R}^N$ (or better $(\mathbb{R}^n)^m$), where $\mathbf{F} = (\mathbf{F}_1, \dots, \mathbf{F}_m)$, $N = nm$ and

$$(14) \quad \frac{d\mathbf{x}}{dt} = \mathbf{F}\mathbf{x}, \quad \mathbf{x} \in \mathbf{S}.$$

Equation 14 is rephrasing Equation 13. This is the genome dynamics of the tissue.

Thus this tissue has a genome dynamics and separates into a individual cell dynamics \mathbf{B}_i for cell i . Let \mathbf{x}_0 be the point of \mathbf{S} defined as $\mathbf{x}_0 = (\mathbf{x}_{0,1}, \mathbf{x}_{0,2}, \dots, \mathbf{x}_{0,m})$, where $\mathbf{x}_{0,i}$ is the equilibrium in \mathbf{B}_i . The weak hardwiring hypothesis asserts that the $\mathbf{x}_{0,i}$ are all the same. Then \mathbf{x}_0 is the equilibrium for genome dynamics for the whole tissue.

The rest of the paper we assume the weak hardwiring for the cells in the tissue.

Extension of monotonicity from the genome dynamics of a cell to the genome dynamics of the tissue

Extension of the definition of monotonicity to many cells is given by:

$$\langle \mathbf{F}\mathbf{x}, (\mathbf{x} - \mathbf{x}_0) \rangle < 0, \quad \mathbf{F} = \prod \mathbf{F}_i, \quad \mathbf{x} \neq \mathbf{x}_0$$

$$\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m) \text{ and } \mathbf{x}_1 = (\mathbf{x}_1^1, \mathbf{x}_1^2, \dots, \mathbf{x}_1^n) \in \mathbf{X}_1, \text{ etc.}$$

Here \mathbf{x}_i^j denotes the amount of j^{th} protein in the i^{th} cell. $\frac{d\mathbf{x}}{dt} = \mathbf{F}\mathbf{x}$ is the dynamics on the basin $\mathbf{B} = \prod \mathbf{B}_i$, $\mathbf{x}_0 = (\mathbf{x}_{0,1}, \mathbf{x}_{0,2}, \dots, \mathbf{x}_{0,m})$ and $\mathbf{x}_0 \in \mathbf{B}$. Observe from weak hardwiring $\mathbf{x}_{0,1} = \mathbf{x}_{0,2} = \dots = \mathbf{x}_{0,m}$.

Example 2: $\mathbf{F}\mathbf{x} = \mathbf{A}(\mathbf{x} - \mathbf{x}_0)$, where $\mathbf{A} = \prod \mathbf{A}_i$, so that \mathbf{A} is a multi-linear map and each $\mathbf{A}_i : \mathbb{R}^m \rightarrow \mathbb{R}^m$ is linear. Then \mathbf{A}_i , for each i is negative definite exactly when this monotonicity holds.

However we are not assuming the linearity of the dynamics. One cannot even get a good model of robust stability of equilibria in the linear setting. One cannot model dynamics with two separate equilibria.

Diffusion dynamics for n proteins: Recall section 4, the diffusion dynamics between cells in a tissue for a protein distribution

$$(15) \quad \frac{d\mathbf{x}}{dt} = -\mathbf{L}\mathbf{x} \text{ or } \frac{d\mathbf{x}_i}{dt} = -\mathbf{L}\mathbf{x}_i \text{ for all } i = 1, \dots, m.$$

Here $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_m)$, $\mathbf{x}_i \in \mathbb{R}^n$ is an n -tuple of proteins or "a distribution of proteins."

Dynamics of a tissue

We will make the hypothesis that the cells described by $\mathbf{F}_i : \mathbf{B}_i \rightarrow \mathbb{R}^m$ have the same dynamics and the \mathbf{B}_i are the same. This is the hardwiring hypothesis. We also suppose that the basins \mathbf{B}_i are convex. These two hypothesis are made so that the diffusion terms in Equation 16 below make sense.

Following the spirit of Turing's paper, we may combine two dynamics (genome dynamics within the cell and diffusion dynamics between cells) into a system (16) that is the object of the study of this paper.

$$(16) \quad \frac{d\mathbf{x}}{dt} = \mathbf{F}\mathbf{x} - \mathbf{L}\mathbf{x}, \quad \mathbf{x} \in \mathbf{B} = \prod_{i=1}^m \mathbf{B}_i$$

We emphasize that differential equation 16 is not necessarily linear in contrast to Turing.

The main Theorem of this paper is:

Theorem 5: The dynamical system $\frac{d\mathbf{x}}{dt} = \mathbf{F}\mathbf{x} - \mathbf{L}\mathbf{x}$ of a tissue (Equation 16) is globally stable with equilibrium \mathbf{x}_0 provided the tissue is hardwired, the basins \mathbf{B}_i are convex and \mathbf{F} satisfies monotonicity.

Lemma 1: If \mathbf{F} is monotone relative to \mathbf{B}, \mathbf{x}_0 , then $(\mathbf{F} - \mathbf{L})$ is monotone.

Proof of Lemma 1. Lemma 1 is true if \mathbf{F} is monotone and if $-\mathbf{L}$ is monotone. First $\langle -\mathbf{L}\mathbf{x}, (\mathbf{x} - \mathbf{x}_0) \rangle \leq 0$ is proved. From weak hardwiring \mathbf{x}_0 is harmonic and so $\mathbf{L}\mathbf{x}_0 = \mathbf{0}$. Moreover $\langle -\mathbf{L}\mathbf{x}, \mathbf{x} \rangle \leq 0$ for any $\mathbf{x} \in \mathbf{S}$, because of $-\mathbf{L}$ is negative semi-definite. See Section 4. Therefore $\langle -\mathbf{L}\mathbf{x}, (\mathbf{x} - \mathbf{x}_0) \rangle \leq 0$. Since $\mathbf{F}\mathbf{x}_0 = \mathbf{0}$, thus it remains only to prove $\langle \mathbf{F}\mathbf{x}, (\mathbf{x} - \mathbf{x}_0) \rangle < 0$ for $\mathbf{x} \neq \mathbf{x}_0$. But \mathbf{F} is monotone by hypothesis. QED.

By Proposition 1 applied to $\mathbf{G} = (\mathbf{F} - \mathbf{L})$ and Lemma 1 we obtain the global stability of equilibrium \mathbf{x}_0 , thus proving Theorem 5.

We will name the property of $\frac{d\mathbf{x}}{dt} = \mathbf{F}\mathbf{x} - \mathbf{L}\mathbf{x}$ in Theorem 5 "emergence."

Theorem 5 establishes that monotonicity implies global convergence of the tissue dynamics to the equilibrium, that is all cells have the same protein distribution, in a strong stable sense ("robustness"). This gives a biological justification for the concept of hardwiring in a tissue. Thus we give a model for "emergence of function."

Remark 3: Explicit solution of $\frac{d\mathbf{x}}{dt} = \mathbf{F}\mathbf{x} - \mathbf{L}\mathbf{x}$ in the linear case

Recall the linear case

$$(17) \quad \frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x} - \mathbf{x}_0) - \mathbf{L}\mathbf{x}, \quad \mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_m).$$

where $(\mathbf{F} - \mathbf{L})$ is not singular. This is in the form

$$(18) \quad \frac{d\mathbf{x}}{dt} = \mathbf{P}\mathbf{x} - \mathbf{Q}.$$

where $\mathbf{P} = \mathbf{F} - \mathbf{L}$ and $\mathbf{Q} = \mathbf{F}\mathbf{x}_0$. Equation 18 has an explicit solution [16]

$$\mathbf{x}(t) = \exp(\mathbf{P}t)\mathbf{C} + \mathbf{P}^{-1}\mathbf{Q}.$$

$\mathbf{x}(0) = \mathbf{C} + \mathbf{P}^{-1}\mathbf{Q}$, therefore $\mathbf{C} = \mathbf{x}(0) - \mathbf{P}^{-1}\mathbf{Q}$. Furthermore, $\lim_{t \rightarrow \infty} \mathbf{x}(t) = \mathbf{P}^{-1}\mathbf{Q}$ if the eigenvalues of \mathbf{P} strictly negative.

Section 6: Turing's paper on morphogenesis

The work of Alan Turing plays an important role in our paper. The main differential equations 16 owe much to [1]. There are some important differences. First we are using nonlinearity for the cell dynamics in contrast to the Turing linear setting. Nonlinearity allows us to address issues of stability, where the second derivative plays a crucial role and we are able to use associated domains of the cell dynamics more in accord with the biology. On the other hand Turing developed his work in a partial differential equations framework, reaction diffusion equations, that reflect a continuum perspective of the nature of cells. That leads to some applications in morphogenesis, such as patterning in Zebra stripes [17, 18, 19, 20]. Our own

perspective differs. We feel that some of the basic features of morphogenesis must deal with few cells (embryogenesis, cell differentiation).

The recent work of Chua [11, 12] also develops Turing's contributions in a different direction from our work.

Turing found an important example of the system of the same type we used in Section 5. The example shows how a system that is stable without diffusion becomes unstable in the presence of diffusion. Turing was motivated to understand morphogenesis with this example of instability. The example consists of two cells and two proteins. The variables x_1, y_1 represent concentrations of molecules (or proteins) for the first cell and x_2, y_2 for the second cell. Turing's two cell reaction-diffusion example can be written as:

$$(19) \quad \begin{aligned} \frac{dx_1}{dt} &= (5x_1 - 6y_1 + 1) + 0.5(x_2 - x_1) \\ \frac{dy_1}{dt} &= (6x_1 - 7y_1 + 1) + 4.5(y_2 - y_1) \\ \frac{dx_2}{dt} &= (5x_2 - 6y_2 + 1) + 0.5(x_1 - x_2) \\ \frac{dy_2}{dt} &= (6x_2 - 7y_2 + 1) + 4.5(y_1 - y_2), \quad x_1, y_1, x_2, y_2 > 0. \end{aligned}$$

The two cells are identical in this example and we can describe the cell dynamics as $\frac{dx_1}{dt} = (5x_1 - 6y_1 + 1)$, $\frac{dy_1}{dt} = (6x_1 - 7y_1 + 1)$. It is easy to transform System 19 into our form $\frac{d\mathbf{x}}{dt} = (\mathbf{F} - \mathbf{L})\mathbf{x}$.

Genome dynamics of the Turing example:

We now show that a key phenomenon of this example is the failure of the monotonicity condition. That is necessary to give rise to instability (morphogenesis).

Let us then study the monotonicity of Section 3 as well as a two dimensional analysis for the Turing example. First we construct matrix \mathbf{A} for a single cell of the Turing example: $a = 5, b = -7, c = 6, d = 7$, and $\mathbf{A} = \begin{pmatrix} 5 & -6 \\ 6 & -7 \end{pmatrix}$. Here the Turing example assumes two identical cells and we can write the monotonicity condition as $\langle (5x - 6y, 6x - 7y), (x, y) \rangle < 0$ for all $x, y > 0$. Thus $5x^2 - 7y^2 < 0$. If $x = 2, y = 1$ we can see that $5x^2 - 7y^2 > 0$; therefore the Turing two cell example fails the monotonicity condition (see the red dot in Figure 4). Monotonicity is only a sufficient condition for stability. Now we check for stability.

The trace(\mathbf{A}) = -2, and det(\mathbf{A}) = 1. Thus, the eigenvalues of \mathbf{A} are given by $\lambda_i(\mathbf{A}) = -1, i = 1, 2$ (and the eigenvectors are given by $v_i = (1, 1), i = 1, 2$). Therefore, the genome dynamics is stable for one cell and hence for two cells.

Diffusion dynamics:

The diffusion dynamics in System 19 is expressed by the terms $0.5(x_2 - x_1)$ and $4.5(y_2 - y_1)$. Since the diffusion dynamics is represented by the negative laplacian matrix as in Section 4, its eigenvalues are non positive.

Full Dynamics:

Combining genome dynamics and diffusion dynamics gives the Turing example of System 19. The linear part of System 19 is

$$\mathbf{M} = \begin{pmatrix} 4.5 & -6 & 0.5 & 0 \\ 6 & -11.5 & 0 & 4.5 \\ 0.5 & 0 & 4.5 & -6 \\ 0 & 4.5 & 6 & -11.5 \end{pmatrix}.$$

System 19 has a unique equilibrium that is obtained by solving the right hand side of the equation set equal to zero. Eigenvalues of \mathbf{M} can be computed to be 2.0, -1 , -1 , -14 . Since there is a positive eigenvalue, the system with diffusion is unstable. Therefore the Turing system not only has a possibility of failure of stability but in fact it is unstable.

Summarizing, the Jacobian at the equilibrium has four eigenvalues, one of which is positive. Thus, the full system with diffusion is not stable. In this way Turing showed that at the equilibrium the two cell example without diffusion is stable, but with diffusion has lost stability. We have remarked further on the role of monotonicity.

Smale [10] examined similar equations with nonlinear cell dynamics. He considered each of two cells as having a global stable equilibrium, and therefore the cells were "dead" in an abstract mathematical sense. But upon coupling the two cells by diffusion, he proved that the resulting system has a global periodic attractor, and hence the cells become "alive."

Towards this end Smale's work was a mathematical model similar to the Turing two cell example but with dynamics of each cell not linear leading to the model,

$$(20) \quad \frac{dx^k}{dt} = R(x^k) + \sum_{i \in \text{set of cells neighboring } k\text{th cell}} \mu_{ik}(x^i - x^k)$$

where $k = 1, \dots, N$ and $(x^i - x^k) \in \mathbb{R}^m$. The first term above $R(x^k)$ gives the dynamics for the k^{th} cell and the second term describes the diffusion processes between cells. The principal case considered by Smale is $m = 4$, $N = 2$, shows for the appropriate choice of parameters (R, μ_k) , the system has stable equilibria without diffusion and with diffusion has a global periodic attractor. The Equation 20 is precisely a form of our main equations. Again the phenomenon depends on the failure of monotonicity

(Easy) Conjecture 1: Generically monotonicity of a linear system on euclidean space is equivalent to all eigenvalues negative (and real).

7. Lapse of emergence

Here we discuss an avenue to study the departure from emergence using our setting. Consider the Jacobian of $\frac{d\mathbf{x}}{dt} = \mathbf{F}\mathbf{x} - \mathbf{L}$ (Equation 16) at the equilibrium \mathbf{x}_0 , that is

$$(\mathbf{D}(\mathbf{F} - \mathbf{L}))_{\mathbf{x}_0}, \quad \mathbf{x}_0 = (\mathbf{x}_{0,1}, \dots, \mathbf{x}_{0,m}),$$

where $\mathbf{x}_{0,i}$ is the equilibrium of the dynamics of the i^{th} cell and $\mathbf{x}_{0,i}$ are all equal. The main cause of lapse of emergence is the vanishing of determinant of $(\mathbf{D}(\mathbf{F} - \mathbf{L}))_{\mathbf{x}_0}$. As in our paper [20], the pitchfork bifurcation is signaled at a bifurcation parameter μ , where the $\det(\mathbf{D}(\mathbf{F} - \mathbf{L}))_{\mathbf{x}_0}$ first becomes zero. Since $\mathbf{x}_{0,i}$ are equal, $\mathbf{L}\mathbf{x}_0 = 0$

and \mathbf{x}_0 belongs to the n -dimensional harmonic space, $\text{kernel}(\mathbf{L})$. Generically $-\mathbf{L}$ is contracting to $\text{kernel}(\mathbf{L})$, that is the eigenvalues of $-\mathbf{L}$ are $\lambda_k = 0$ for $k \leq n$, and for $k > n$, $\lambda_k < 0$.

Now look at $\mathbf{D}(\mathbf{F})_{\mathbf{x}_0} = \left(\mathbf{D}(\mathbf{F}_1)_{\mathbf{x}_{0,1}}, \mathbf{D}(\mathbf{F}_2)_{\mathbf{x}_{0,2}}, \dots, \mathbf{D}(\mathbf{F}_m)_{\mathbf{x}_{0,m}} \right)$ on the basin \mathbf{B} . Each $\mathbf{D}(\mathbf{F}_i)_{\mathbf{x}_{0,i}}$ is contracting before the bifurcation, say for $\mu < 0$. At $\mu = 0$, one can expect one of these contracting derivatives to become singular, for example $\mathbf{D}(\mathbf{F}_1)_{\mathbf{x}_{0,1}}$ with one eigenvalue equal to zero and the rest are negative. This is the beginning of the lapse of emergence in this scenario. Now restrict the dynamics to the protein space of the first cell to study the bifurcation. In this protein space we can expect the dynamics after the bifurcation to have two basins. This is the setting of the pitchfork bifurcation paper [21]. This paper can be used to examine the end of emergence in terms of cell division (symmetric or asymmetric, or cancer) [22, 23].

REFERENCES

- [1] Turing A (1952) The chemical basis of morphogenesis. *Philos. Trans. Roy. Soc. London, Series B* 237: 37-72.
- [2] Rajapakse I, Smale S (2016) Mathematics of the genome. *Foundations of Computational Mathematics* 1-23.
- [3] Browder F (1983) Degree of mapping for nonlinear mappings of monotone type. *Proc Natl Acad Sci USA* 80 (6): 1771-1773
- [4] Browder F (1983) Degree of mapping for nonlinear mappings of monotone type: Densely defined mapping. *Proc Natl Acad Sci USA* 80 (8): 2405-2407.
- [5] Hirsch M, Smith H (2003) Competitive and Cooperative Systems: A Mini-review, *Proceedings of the First Multidisciplinary International Symposium on Positive Systems: Theory and Applications* 183-190.
- [6] Strang G (2003) Introduction to Linear Algebra, Wellesley-Cambridge Press.
- [7] Strogatz S (2000) Non-linear Dynamics and Chaos: With applications to Physics, Biology, Chemistry and Engineering (Perseus Books).
- [8] Cavalli G, Misteli T (2013) Functional implications of genome topology. *Nature Structural & Molecular Biology* 20(3): 290-299.
- [9] Rajapakse I, Groudine M (2011) On emerging nuclear order. *Journal of Cell Biology* 192(5): 711-721.
- [10] Smale S (1974) A mathematical model of two cells via Turing's equation. *Lectures in Applied Mathematics*, Vol. 6, (American Mathematical Society), 15-26.
- [11] Chua L (2005) Local activity is the origin of complexity. *International Journal of Bifurcation and Chaos* 15 (11): 3435-3456.
- [12] Mainzer K, Chua L (2013) Local Activity Principle (Imperial College Press. London).
- [13] Rauch E and Millonas M (2004) The role of trans-membrane signal transduction in turing-type cellular pattern formation. *Journal of Theoretical Biology* 226: 401-407
- [14] Chung FR (1997) Spectral Graph Theory (American Mathematical Society, Providence, RI).
- [15] Cucker F, Smale S (2007) Emergent Behavior in Flocks. *Automatic Control, IEEE Transactions* 52: 852-862.
- [16] Brockett RW (1987) Finite Dimensional Linear Systems. John Wiley & Sons, Inc., New York, USA.
- [17] Kondo S, Miura T (2010) Reaction-diffusion model as a framework for understanding biological pattern formation. *Science* 329 (5999): 1616-620.
- [18] Maini P, Baker R, Chuong C (2006) Developmental biology. The Turing model comes of molecular age. *Science* 314(5804): 1397-1398.
- [19] Murray JD (2002) Mathematical Biology, third ed., Springer-Verlag.
- [20] Tompkins N et. al (2014) Testing Turing's theory of morphogenesis in chemical cells. *Proc Natl Acad Sci USA* 111: 4397-4402
- [21] Rajapakse I, Smale S (2016) The Pitchfork Bifurcation. arXiv:1609.05996.

- [21] Misteli T (2009) Self-organization in the genome. *Proc Natl Acad Sci USA* 106: 6885-6886.
- [23] Hartwell L et al (1989) From molecular to modular cell biology. *Nature* 402 C47-C5.12

UNIVERSITY OF MICHIGAN

E-mail address: `indikar@umich.edu`

UNIVERSITY OF CALIFORNIA BERKELEY

E-mail address: `smale@math.berkeley.edu`